DOCUMENT RESUME

ED 248 790 HE 017 655

AUTHOR Boli, John; And Others
TITLE Grades, Grading Standards, and Academic Awards. Fourth in a Series about the Stanford Curriculum Study.
INSTITUTION Stanford Univ., CA. Office of Undergraduate Research.
SPONS AGENCY Metropolitan Life Foundation.
PUB DATE [84]
NOTE 35p.; For related documents, see HE 107 652-654.
PUB TYPE Reports - Research/Technical (143)

EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS *Academic Achievement; Aptitude Tests; *Awards; Class Size; *College Students; Comparative Analysis; *Departments; Engineering; *Grades (Scholastic); Grading; Higher Education; Humanities; Institutional Research; Scores; Trend Analysis
IDENTIFIERS Stanford Curriculum Study; *Stanford University CA

ABSTRACT

Grading standards in departments and school divisions at Stanford University were analyzed, along with grading trends, and the classroom performance of students with different majors. From 1975-1976 to 1981-1982, the proportion of A grades decreased, while B and C grades increased. In 1976, D grades were reinstituted. Considerable variation in grading standards of the various school divisions was found. Grading was highest in the humanities, with nearly half of all grades reported as A's, while it was lowest in the School of Engineering. Even more variation in grading standards was found for individual departments. Generally, the level of grades varied inversely with the degree of technical content of the subject area, with the national sciences and engineering having the lowest grade distributions and the humanities and social sciences the highest. Grades were inversely related to class size in all of the school divisions. Humanities classes had greater enrollments in smaller classes, while the natural sciences had greater enrollments in large classes. Natural science majors did better in nearly all fields and were more likely than other students to earn awards at graduation. Information is also provided on the relationship of grades and Scholastic Aptitude Test scores. (SW)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED248790

GRADES, GRADING STANDARDS, AND ACADEMIC AWARDS
Fourth in a Series About the Stanford Curriculum Study

John Boli
Research Coordinator
Office of Undergraduate Research

Herant Katchadourian
Vice Provost for Undergraduate Education

Sally Mahoney
Registrar

Stanford University

U.S. DEPARTMENT OF EDUCATION
NATIONAL INSTITUTE OF EDUCATION
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☑ This document has been reproduced as received from the person or organization originating it.
☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official NIE position or policy.

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

John Boli

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

HE 017 555

We would like to thank Jim Sharp for his programming and analytical work, Nancy Olsen for her very helpful readings of these articles, and Andrea Garwood for producing the many drafts of the articles and tables. We are also grateful to Lew Mayhew, Hans Weiler, and Peter Kahn for their critique of earlier versions of this material, and to the Metropolitan Life Foundation for funding part of this study.

GRADES, GRADING STANDARDS, AND ACADEMIC AWARDS

ABSTRACT

An examination of grades earned by undergraduates is a major element of the Stanford Curriculum Study. This article analyzes grading standards in departments and school divisions, grading trends over time, and the classroom performance of students of different majors. It also includes multiple regression analyses evaluating the effects of various student and course characteristics on grades earned. Policy implications are discussed.

GRADES, GRADING STANDARDS, AND ACADEMIC AWARDS

Fourth in a Series About the Stanford Curriculum Study

This series of articles describes the results of the Stanford Curriculum Study, a longitudinal examination of the background characteristics and academic records of graduating seniors. The study is used to inform policy-making bodies concerned with undergraduate education at Stanford, and these articles illustrate the kinds of analysis carried out for the study and the uses to which the results of the study can be put.

In the three previous articles in the series we discussed degrees granted, course enrollments, and sex differences in academic work among Stanford students [2, 3, 4]. In this final article we turn to the issue of grades. We examine the grades given in courses in the various school divisions or departments (hence, the grading standards of academic units) as well as the grades earned by students having different majors or belonging to different groups of majors (hence, the performance of students of different majors). Our interest in analyzing grades is part of a larger effort to monitor the quality and characteristics of the undergraduate program. The issues we study include the following:

(a) Is there any evidence of a trend in the grading standards at Stanford over the seven years covered by the Curriculum Study?

(b) How do grading standards vary across major school divisions and departments? Where are grades highest and where are they lowest? Are these differences the result of different standards, or are they artifactual products of structural differences (class size, instructor type, and so on) among school divisions?

(c) Are there systematic differences in the grades earned by students in their own major departments compared with those earned by students outside their own majors?

(d) How much variation is there in the classroom performance of students of different majors or school divisions? Do the students of some majors receive better grades and more academic awards than other students? What accounts for these performance differences?

(e) Finally, how are grades at Stanford related to student background characteristics and course or departmental characteristics? In other words, to what extent can the grades of Stanford degree recipients be predicted?

METHODOLOGICAL CONSIDERATIONS

The grading system at Stanford consists of letter grades "A" through "D" and "pass." There is no "F" grade. Students who fail a course simply do not receive credit for the course and must take it again if it is required for their program of study. Students can choose to take courses on a "pass/no credit" basis, where the "pass" notation is awarded for minimally satisfactory (or better) performance. A number of courses are offered only on a pass/no credit basis, i.e., no letter grades are awarded. About half of the pass grades are by student choice and half are at the option of the instructor.

Because the grading system differs from that of most schools in its treatment of students failing courses, Stanford does not compute grade point averages (GPAs). We have done so here in order to facilitate comparisons, using a scale ranging from 1.0 (D) to 4.0 (A). We take "+" and "-" notations into account by assigning values in units of 0.3 (e.g., "A+" is assigned 4.3, "B-" is assigned 2.7). A GPA greater than 4.0 is thus possible, though extremely rare. In the analyses reported here we omitted "pass" grades in calculating means, but in parallel analyses we assigned them the value of 2.0 and found that the results did not differ substantially for the two methods.

Many of our comparisons use the "school division" as the unit of analysis. A school division is a group of related departments and programs;

at Stanford there are six, including four divisions of the School of Humanities and Sciences (the humanities, social sciences, natural sciences, and interdepartmental programs) and the Schools of Engineering and Earth Sciences. The school divisions contain between four and fifteen departments and programs each.

Finally, there is great stability in the grades given in the various courses and departments of the university, as we shall show below, and relatively few systematic trends in the data. We therefore will present year-by-year results only in a few instances, relying instead on mean figures or detailed analyses of a single year's data to investigate special topics. In the latter case we usually use data for the graduates in 1979-80 because this year was near the middle of the seven-year period we have studied and was quite typical of the entire body of data.

GRADING STANDARDS OVER TIME

The issue of whether there has been a change in grading standards at Stanford over the period of our study is examined in Table 1. The table shows the distribution of all grades for each graduating cohort. Put simply, the stability of these data is remarkable. Year-by-year variability in the percentage of grades in any of the five categories was no more than two percentage points. But there was also a very gradual declining trend: the proportion of A's was decreasing while B's and C's were increasing and the use of the pass/no credit option was waning. Hence the well-known grade inflation of earlier times [8] was reversed in recent years, at least at Stanford.

It is important to note that the distribution of class size was very stable over this time period. The importance of this fact is as follows. Suppose that there had been a trend toward larger classes in the university as a whole. Given a negative relationship between class size and grading (see

below), the modest decline in grades would reflect nothing more than the change in average class size. However, when class size is very stable and grades decline slightly, as in our case, we can infer that the deflation in grades was real rather than only apparent.

One possible explanation for the grade deflation evident in Table 1 is the resurrection of the 'D' grade at Stanford in 1976-77. Up until then there had been no 'D' for a number of years; the concern over grade inflation led to its reinstitution. While the overall grade distribution was almost identical from 1975-76 to 1976-77, in the longer term instructors became more willing to give 'B' and 'C' grades even though they never gave more than one percent 'D' grades. On the other hand, the decline in 'pass' grades after the 'D' was revived is contrary to expectation, given that the lowest possible passing letter grade was now a step lower.

One of the key policy issues raised by Table 1 is the appropriateness of the general grading standards at Stanford. Considering the very slight stiffening of grading practices in recent years, there is little need to be concerned about trends in those standards. But how does the university decide if grading is "too high" or not? On the one hand, it has to be admitted that grades are very high when the average GPA is above 3.4, or 'B+'. On the other hand, and this is the theory that seems to have prevailed at the university, Stanford students are all of very high caliber and their "average" performance is very good. Relative to students at most other institutions, Stanford students "ought" to produce high-quality academic work. The high grades given to the students thus simply reflect the ability of the students and are not unreasonably high (cf. Astin's [1] finding that grading is considerably higher in highly selective colleges than in other schools). No systematic attempt to go beyond such theorizing has been made (e.g., comparing grading standards at

Stanford with those at other schools), and generally it is fair to say that university policy-makers have shied away from a serious consideration of the appropriateness of grading standards in undergraduate courses.

GRADES GIVEN BY SCHOOLS AND DEPARTMENTS

There is considerable variation in the grading standards of the various school divisions [cf. 10]. In Figure 1 we present both the percentage distributions and the mean grades by school division. The figures are based on only five years of the Curriculum Study period (for technical reasons data are not available for the other two years). Grading was highest in the humanities, with nearly half of all grades reported as "A's," while it was lowest in the School of Engineering; the difference in average grades on a 4-point scale was about 0.30. Further, grading in the humanities was well above all other school divisions.

When we turn to individual departments we find even more variation in grading standards. In Table 2 we give the mean grades by department for 1979-80 graduates; these figures were typical of the entire period. Mean grades ranged from a high of 3.80 in music courses to a low of 2.97 in chemical engineering courses. Within school divisions we find even more variation than across the divisions; in engineering, for example, the range was fully 0.38 units (chemical versus electrical engineering) and in the humanities the range was as high as 0.53 (music versus history). The most uniform grading occurred in the social and natural sciences, but here too there were clear differences from one department to another.

Generally speaking, the level of grades given varied inversely with the degree of technical content of the subject area, with the natural sciences and engineering having the lowest grade distributions and the humanities and social sciences the highest.[1] Even within school divisions this was the

case: for example, the most technical social science discipline, economics, had the lowest grades; the most empirically-oriented humanities discipline, history, had the lowest grades in its division.

Are the Grading Differences Real?

While the more technical subject areas had the lowest grades, they also had generally larger classes. The humanities had the highest proportion of enrollments in smaller classes (31% in classes of 15 or fewer students), followed by the earth sciences (15%) and social sciences (10%). At the other end of the scale, the natural sciences had the highest proportion of enrollments in large clasres (47% in classes of more than 120 students) and the humanities the lowest (only 5%). Engineering had relatively few enrollments in large classes (24%) but a very high proportion in moderately large classes (57% of enrollments in classes of 30-120 students). What influence does class size have on grading at Stanford?

This question is answered in Figure 2. Here we present the mean grades for courses of different class sizes[2] in each of the four largest school divisions. The figure shows, for example, that the mean grade for engineering classes of one to ten students was 3.70, falling to only 3.06 for engineering classes of 257 or more students. Grades were inversely related to class size in all of the school divisions. The relationship was strongest in the natural sciences and engineering and quite substantial in the humanities and social sciences. In no case was it perfectly monotonic; for example, there was some increase in grading for the very largest classes compared to the next largest category, but overall the relationship was very clear.

Figure 2 shows that the differences in grading across the school divisions were very greatly attentuated when class size was controlled. Engineering, the division having the lowest overall grades, had only slightly

lower grades than the humanities at each of the class sizes up to the category of 23 to 32 students, and in all categories the difference between these two divisions was less than the overall mean grade difference of 0.24. Grades in the natural sciences also differed only moderately from those in the humanities and social sciences when class size was controlled. Indeed, for the smallest class size the grades in engineering were as high as those in the humanities and social sciences and grading was higher in natural science courses than in other courses.

Our general conclusion is that, while the humanities and social sciences had higher grades than the other divisions, the differences were considerably exaggerated by the smaller size of classes in these divisions. Further support for this finding is provided by calculating the mean 1979-80 grades that would have been obtained if class size in each of the four largest school divisions were distributed as in the humanities. These results are quite revealing:

Humanities	3.45
Social sciences	3.43
Natural sciences	3.42
Engineering	3.33

Here we see that there are virtually no differences among the humanities, natural sciences, and social sciences. Engineering remains as the school with the lowest average grades, but now its mean lies only about 0.12 points below that of the humanities rather than, as in the unadjusted data, fully 0.24 points below.[3]

We have also examined the relationship between grading and type of instructor, finding that the average grades given by regular professorial faculty (full, associate, and assistant professors) varied little across the

three ranks but were lower than the grades given by other types of instructors (lecturers, teaching fellows, etc.). Overall the grades given by professorial faculty averaged about 3.32, while for other instructors the average was about 3.45 (1979-80 data). Breaking down the mean grades for instructor type within each school division showed that the professorial faculty in the humanities gave consistently higher grades than their counterparts in other divisions, including the social sciences. However, because the humanities had the lowest percentage of course enrollments by professorial faculty (52%, compared to 70% and 71%, respectively, in engineering and the natural sciences), controlling for instructor type does not reduce the difference in mean grades between humanities and the other school divisions as much as controlling for class size.

The overall result of this analysis is that the grading differences among school divisions are less than initially meet the eye. To a large extent, grades are higher in the humanities than in engineering because classes in the former are smaller than those in the latter. This is an important result that should help moderate the view that grading in the humanities is "easy," and grading in engineering "tough." The idea that natural science courses are more rigorously graded is also shown to be erroneous; there is little difference in mean grades among the natural sciences, social sciences, and humanities when class size differences are taken into account.

GRADES EARNED BY STUDENTS OF THE VARIOUS SCHOOL DIVISIONS

We have seen that, even with controls for class size and instructor type, grading was somewhat lower in engineering than in the other school divisions. One possible explanation for this finding is that engineering students simply performed less well than students majoring in other areas (the majority of enrollments in engineering are taken by the school's own students

rather than by students from other divisions). This hypothesis leads us to an examination of grades earned by students majoring in each school division in the courses they have taken in the various subject areas. If we find that engineering students performed less well than other students in all course areas of the university, the lower grading in engineering courses may be attributed to the weaker ability of engineering students. If, on the other hand, engineering students performed as well in courses both within and outside engineering as other students, we are brought closer to the conclusion that grading standards in engineering courses really were somewhat stricter than standards in other courses.

Table 3 presents the mean grades earned by student majors of each school division in the courses of each division (1979-80 data). We omit the earth science and interdepartmental courses because they were relatively few in number. Beginning with courses in engineering, we see that engineering students earned slightly lower grades than students in three other school divisions, and very slightly better grades than students in two divisions. Their grades were also lower than average in humanities courses and natural science courses, while they were slightly higher than average in social science courses. On the basis of these findings we cannot conclude that engineering students performed less well in general than other students. Even where they did less well than other students, the differences were not large.

We conclude that grading was indeed somewhat lower in engineering than in other areas. This view is further supported by the fact that the grades earned in engineering were lower overall than grades earned in any other of the four school divisions.

Table 3 also shows that grading in natural science courses was lower than in other areas except engineering. At the same time, natural science students

performed better than all other groups of majors across the entire range of courses in the university: they had the highest average grades in each of the four school divisions. But the grades earned in natural science courses were lower than those earned in any other area for earth science, social science, and interdepartmental majors, and they were near the bottom of the ranking for engineering and natural science majors. Hence, the lower grades in natural science courses were clearly not the result of poorer performance by the division's own majors; most groups of students did less well in the natural sciences than they did elsewhere.

In addition to the grading standards of the school divisions, the relative performance of students grouped by major school division can be evaluated in Table 3. As mentioned above, the natural science majors received higher grades than any other group of majors in the courses of all four school divisions. There has clearly been a tendency for pre-medical students to achieve stronger academic records than other students.[4] Of the other major groups, none of them stands out as having done particularly well or poorly, with the possible exception of the earth science students. Some of the findings are surprising; for example, humanities majors had the second highest grades in natural science courses (and a not insubstantial number of enrollments in that division), while earth science students received lower grades than nearly all other groups in both engineering and natural science courses. It is also somewhat surprising that students majoring in a given division were by no means certain to earn the highest grades in that division. In only one of the four largest divisions was this the case (natural sciences), though humanities students did better than all but natural science students in humanities courses. Otherwise, we find that engineering students performed less well than three other major groups in engineering

courses, and social science students received only average grades in social science courses.

In sum, natural science students distinguished themselves in classroom performance while differences among the other major groups were more mixed--although there is some evidence that earth science students tended to receive lower grades than the other groups.

Should university policy-making bodies be concerned about the differences in performance of students in the various school divisions? At Stanford this question has received a conditionally affirmative answer. The concern is not with the tendency of more able students to concentrate in the natural sciences but with the consequences of that concentration - the "pre-med syndrome" of extreme emphasis on grades rather than learning, intense competition, and an unwillingness on the part of students to take courses outside the pre-medical track for fear of receiving lower grades or "wasting" units on "irrelevant" courses. The university has not tried to discourage high-ability students from entering the natural sciences, but it has tried to encourage those students to seek a broad education and not limit themselves to natural science areas. On the whole we believe that these efforts have been successful. Actually reducing the competitiveness of the pre-medical students is a more intractable problem, and there has been little effort in this direction.

AWARDS EARNED AT GRADUATION: PHI BETA KAPPA AND DISTINCTION

Two academic awards are highly dependent on grades earned by students: Phi Beta Kappa and University Distinction. Other criteria enter into the selection process, but grades play the largest role. Election to Phi Beta Kappa is limited to ten percent of any graduating class and is viewed as the most prestigious academic honor a student can earn. University Distinction is solely a Stanford award and can be given to at most 25% of the graduates in

any class.

Table 4 shows the percentages of students in each school division receiving Phi Beta Kappa and Distinction. As can be seen in the last row of each section of the table, the university means were ten percent earning Phi Beta Kappa and 23% earning Distinction, with the former declining by three percentage points over the six years and the latter declining by at least two points. These declines reflected the judgment of the faculty members who make these awards that fewer students were pursuing a sufficiently broad undergraduate education with truly outstanding performance to merit them.

We saw in the previous section that natural science students performed best academically in all areas of the university. Table 4 further supports this finding, as the percentage of students earning Phi Beta Kappa was highest for the natural science students. It was also higher than average for humanities students. Engineering and earth science students were least likely to receive this award. This result was probably due more to the relatively narrow range of courses taken by engineering majors (due to the requirements of their major programs) than to lower academic ability of these students.

Turning to the figures for University Distinction, humanities and natural science students were most likely to earn this honor while earth science students were again least likely. The decline in the percentage of students awarded Distinction was evident almost across the board but it was most severe for the interdepartmental majors.

For both of these awards there was a good deal of variation among the individual departments within the several schools. Large departments that had high percentages of students earning Phi Beta Kappa include history, international relations, and chemistry, while departments in which students were most likely to earn Distinction include the latter two and art, music,

chemical engineering and, well above all other departments, physics. On the other hand, very few students of some departments earned these awards (e.g., less than five percent of majors in communication and several engineering and earth science departments earned Phi Beta Kappa, and less than 15% of majors in communication, geology, and human language received Distinction). Variations of this sort were quite stable, implying that some majors consistently attracted academically more able students than did others. But we should temper this conclusion with the caution that departmental standards for recommending students for Distinction vary widely. The award cannot be interpreted solely as an indicator of students' academic performance.

GRADES AND SAT SCORES: CAN GRADES BE PREDICTED?

As a highly selective undergraduate institution, Stanford has a student population of very high demonstrated scholastic ability. There is, nevertheless, considerable variation in college entrance test scores (in particular, Scholastic Aptitude Test scores) and other variables that might affect college classroom performance. In this section we consider the relationship between SAT scores and grades, performing analyses that use SAT scores and a number of other independent variables to show the extent to which grades can be "predicted" (cf. [7, 9]).

We begin with zero-order correlations between SAT scores and grades, as given in Table 5 (data for 1979-80 graduates only). We present the correlations for verbal and mathematics scores separately, giving figures for grades in all courses, courses in the larger school divisions, and courses taken by each division's group of graduating students. None of the correlations was of very large magnitude, with the highest correlations (0.24) between the SAT mathematics scores and grades in engineering and natural science courses. However, all of the coefficients were statistically

significant beyond the .001 level due to the large number of cases.

Of greater interest in the table is the pattern of relationships. The correlations between grades in natural science and engineering courses and SAT mathematics scores were substantially higher than the correlations between these grades and SAT verbal scores. For social science courses, which call on more of a blend of verbal and mathematical abilities, the correlations of grades with the two test scores were nearly identical (and are among the larger figures in the table). And for humanities courses, verbal scores were somewhat more highly correlated with grades than were math scores.

This pattern also generally characterizes the grades earned by students majoring in the various school divisions, with the important exception of humanities students. For the latter, the verbal coefficient was only very slightly larger than the mathematics coefficient. Interestingly enough, the relationship for interdepartmental majors was larger for the verbal score than the mathematical score, indicating that the mathematical ability of these students was less crucial for successful classroom work than their verbal ability even though most of them were in the human biology program.

We have also examined the correlations for courses in individual departments having a thousand or more enrollments. The largest correlation obtained here was 0.30, between grades in chemistry courses and the SAT mathematics score (for mathematics courses the corresponding figure was 0.24); the largest correlation between the SAT verbal score and grades came in economics (0.27), and economics also showed the largest pair of correlations (the mathematics score was correlated 0.28 with grades in economics). The pattern of relationships generally followed our expectations about the various disciplines with only a few exceptions. Furthermore, almost all of the correlations were significant beyond the .001 level, and the only department

for which both correlations were not significant at even the .05 level was computer science. In these courses grades were almost entirely unrelated to SAT scores. In sum, grades were somewhat related to SAT scores but the relationships varied in sometimes unexpected ways.

In order to determine the degree of predictability of grades at Stanford, we carried out a series of multiple regressions using grades as the dependent variable. As independent variables we used the following characteristics of students and courses:

(a) SAT verbal and mathematics scores (raw scores);

(b) Sex of student, coded as a dummy variable (women coded 0, men coded 1);

(c) Class size, using the deciles of Figure 2 coded from 1 (1-8 students) to 10 (256+ students).

(d) Instructor type, coded as a dummy variable (professorial faculty coded 0, all other instructors coded 1);

(e) Course level, coded as a dummy variable (course number less than 100 coded 0, course number greater than 100 coded 1--a rough way to distinguish introductory courses from advanced courses);

(f) "Own school division," a dummy variable assigning a score of 1 for courses in the student's own school division and 0 for courses outside the school division.

(g) Science and technology courses, a dummy variable coded 0 for natural science and engineering courses and 1 for other courses.

All of the dummy variables were coded such that the cases receiving a score of 0 were those for which lower grades were expected (e.g., in general we expected lower grades in introductory courses than in advanced courses and lower grades for courses outside the student's own school division than within

it). Hence, for each of the dummy variables (except sex) we expected a positive effect on grades received. For the SAT variables we expected positive relationships, while for sex we expected differing relationships depending on the types of courses involved but no overall sex-related effect. For class size we expected a negative effect (cf. Figure 2).

Table 6 presents the results of the multiple regression for 1979-80 data, with all variables entering the equations simultaneously. The coefficients are presented for all enrollments (excluding physical education courses) and for enrollments in the courses of five school divisions. For the equation using all enrollments the table presents both unstandardized (b) coefficients and standardized (Beta) coefficients, as well as the zero-order correlations, standard errors, and F-ratios for each variable. For the other equations we give only the Beta coefficients. We also supply the r-square and constant for each equation.

For all cases we find results completely in line with our expectations: both SAT verbal and mathematical scores had positive effects on grades, while class size had a negative effect and each of the dummy variables (except sex of student) had a positive effect. The largest effect was that of the mathematics SAT score, followed closely by class size, while the three other "characteristics of the course" variables had substantively very small effects. In terms of the specific effects of particular variables, it might be helpful to interpret the b-coefficients in literal terms. Thus, the coefficient for the SAT verbal score (0.00088) indicates that a student with a score 100 points higher than another student can be expected to have a grade point average about 0.09 units higher. Similarly, grades in introductory courses can be expected to be about 0.03 units lower than those in advanced courses (a practically negligible amount from a substantive point of view).

Hence, each of the variables in the equation had discernible effects on grades received. However, the overall ability of the equation to predict grades was quite limited; the multiple r-square of 0.086 indicates that all of the variables in combination accounted for only about nine percent of the variance in grades.

What about the effects for the courses in each school division? Here we have perhaps more interesting results. Let us consider each of the independent variables in turn. The SAT scores had varying effects, with the verbal score most affecting grades in the humanities and social sciences and the mathematics score most affecting grades in engineering and the natural sciences. These results are consistent with our expectations. The sex of the student had effects only in the areas of the earth sciences and natural sciences, and they were in opposing directions: women performed better than men in earth science courses but men performed better in natural science courses. In other areas, including engineering, there were no significant sex-related effects when the other variables were taken into account.

Class size had the strongest effects of all the course characteristics. Its coefficients were largest for earth science and humanities courses but smaller for other courses. We have no substantive explanation for this finding.

Other course characteristic variables had uniformly small effects that were mostly consistent across the divisions. Majoring in the school in which the course was offered was most advantageous for natural science majors, but this finding was probably a reflection of the higher general academic performance of natural science majors. In social science and engineering courses this relationship was reversed; majors from outside these school divisions did better than the divisions' own majors. The instructor type

variable also had its strongest effects in natural science courses, indicating that it was in that division that professorial faculty differed most from non-professional faculty in their grading practices; professional faculty applied stricter standards. Course level, finally, had small positive effects overall -- grades were higher in advanced courses -- but in the earth sciences a small negative effect was found. This can be explained by the fact that earth science students received slightly lower grades in earth science courses than students from outside the school.

Once again, nearly all of the coefficients in the equations for the school divisions were significant; but our ability to predict grades was quite limited in all instances, with no r-square higher than 0.102. Thus, while a number of variables had important effects on grades earned, most of the variability in grades at Stanford could not be accounted for by the quantitative indicators available to us.[5]

But the findings in these analyses are important for policy concerns nonetheless. First, it is encouraging to find that, controlling for SAT scores and other variables, women's grades were not lower than men's grades in engineering courses and were only slightly lower in natural science courses. This finding indicates that women are not particularly impeded by the traditional view that technical fields are not "appropriate" for women. Though we have not studied the issue in depth, it is reasonable to believe that university efforts to encourage women in science and engineering have contributed to women's relatively strong performance in these areas, and these programs should be continued.

Second, the large effect of class size throughout the analyses may be cause for reflection. Is the performance of students really better in smaller classes, or is it simply more difficult psychologically for instructors to

give lower grades to students in small classes? While the answer to this question can be provided only by a separate study, it is worth asking whether the university should be concerned if students are receiving better grades in smaller classes solely because they are small. At the very least, this finding should be taken into consideration in making academic awards at graduation. If a student has had relatively few classes smaller than 30 students, his grades are likely to be lower, through no fault of his own, than those of a student who has had many small classes. On the other hand, we think it would be unwise to try to induce instructors to stiffen their grading standards in smaller classes, for we are well aware of the great educational advantages of small classes and do not want to imply that the higher grades given in small classes are usually unjustified.

Conclusion

The overall grading standards at Stanford were quite stable during the years covered by the Curriculum Study, but grading varied considerably among the six school divisions and even more among the individual departments. Much of this variation could be explained by differences in class size. Even accounting for class size, however, grading was more stringent in engineering than in other school divisions. Engineering students, on the other hand, were found to perform about as well as most other students, and the only group whose classroom performance was significantly different from other groups was the natural science majors. They did better in nearly all fields than other students, a clear indication that the pre-medical subculture attracts high-achieving students. The natural science students were also more likely than other students to earn awards at graduation.

Our multiple regression analyses of grades earned found effects of only small substantive significance, and rather little of the variance in grades

earned could be accounted for by our variables. This is not very surprising, given that most students at Stanford have strong academic abilities and that idiosyncratic, unmeasured factors (individual motivation, interest in course content, etc.) probably account for much of this variance. On the other hand, some of the results of these analyses have important policy implications, especially the very small effect of student sex and the larger effect of class size.

This article concludes our series on the Stanford Curriculum Study. We want to emphasize once again the many benefits to be gained from the analysis of student academic records, and we hope that these articles are useful to other institutions that would like to put their policy-making processes on a firmer foundation by conducting systematic research into student behavior. We will be most happy to correspond with interested readers about the Stanford project and to provide them with a copy of the annual Curriculum Report upon request.

FOOTNOTES

1. The same result was found in a study of grading standards at the graduate level by Prather and Kodras [11], and by Goldman et al. [5] for undergraduates.
2. While the data in Figure 2 concern only 1979-80 graduates, the extreme stability of grading data assure us that similar results would be obtained for any other year.
3. This is not to say that grading in engineering courses necessarily would rise if class size were reduced in that school. Based on the evidence that we can bring to bear on the issue, however, it is safe to conclude that similar class size distributions in the school divisions would considerably attenuate the differences that currently exist.
4. Nearly half of all natural science majors applied to medical school each year.
5. We also examined the predictability of grades in a number of individual coures, heeding Goldman and Slaughter's [6] warning that variable grading standards in different fields can disturb the results of analyses using courses from many academic areas. The results were generally parallel to those in Table 6, with somewhat higher r-squares (as high as 0.21) in some courses. Overall there was little better predictability of grades in individual courses than in groups of courses.

REFERENCES

1. Astin, A. W. Predicting Academic Performance in College. New York: Free Press, 1971.
2. Boli, J., H. Katchadourian, and S. Mahoney. "Analyzing Academic Records for Informed Administration: The Stanford Curriculum Study." 1983a. Journal, pp. to be determined.
3. _____. "Degrees Granted and Course Enrollments by Field of Study." 1983b. Journal, pp. to be determined.
4. _____. "Sex Differences in Student Characteristics and Performance." 1983c. Journal, pp. to be determined.
5. Goldman, R. D., D. E. Schmidt, B. N. Hewitt, and R. Fisher. "Grading Practices in Different Major Fields." American Educational Research Journal, (Fall 1974), 343-357.
6. Goldman, R. D., and R. E. Slaughter. "Why College Grade Point Average is Difficult to Predict." Journal of Educational Psychology, 68 (1976), 9-14.
7. Lavin, D. E. The Prediction of Academic Performance. New York: Russell Sage Foundation, 1965.
8. Levine, A., and J. Weingart. Reform of Undergraduate Education. San Francisco. Jossey-Bass, 1973.
9. Loyd, B. H., R. Forsyth, and H. D. Hoover. "Relationship of Elementary and Secondary School Achievement Test Scores to Later Academic Success." Educational and Psychological Measurement, 40 (Winter 1980), 1117-1124.
10. Prather, J. E. "The Relationship of Major Field of Study with Undergraduate Course Grades: A Multivariate Analysis Controlling for Academic and Personal Characteristics and Longitudinal Trends." ERIC

document ED 132-898. Atlanta: Office of Institutional Planning, Georgia State University, August, 1976.

11. Prather, J. E., and J. E. Kodras. "Graduate Students' Grades for 55 Major Fields: A Multivariate Study." ERIC document ED 150-908. Atlanta: Office of Institutional Planning Georgia State University, October, 1977.

TABLE 1

DISTRIBUTION OF GRADES, ALL COURSES,
1975-76 TO 1981-82
(Percentages)

Grades Given	1975-76*	1976-77	1977-78	1978-79	1979-80	1980-81	1981-82	6-Year Mean**
A	45	44	44	42	43	42	42	43
B	30	30	31	32	31	33	32	31
C	7	6	7	7	7	7	9	7
D	-	1	1	1	1	1	1	1
Pass	19	19	18	18	18	17	16	18

*Grading scale did not include "D" grades in 1975-76.
**1975-76 excluded because grading scale differed from later years.

TABLE 2

MEAN GRADES GIVEN BY DEPARTMENT OR PROGRAM,
1979-80 GRADUATES*

	Mean Grade	Enrollments
Humanities		
Art	3.48	1882
Classics	3.57	584
Drama	3.77	808
English	3.43	4058
French/Italian	3.41	1180
German Studies	3.51	808
History	3.27	2275
Linguistics	3.61	292
Music	3.80	1957
Philosophy	3.43	899
Religious Studies	3.37	345
Spanish/Portuguese	3.66	809
Social Sciences		
Anthropology	3.41	972
Communication	3.39	986
Economics	3.21	3354
Political Science	3.33	2712
Psychology	3.45	3127
Sociology	3.50	874
Natural Sciences		
Biology	3.33	2326
Chemistry	3.12	3435
Computer Science	3.70	1040
Mathematics	3.17	3582
Physics	3.35	2840
Statistics	3.24	490
Interdepartmental		
Human Biology	3.19	2495
Engineering		
Chemical	2.97	394
Civil	3.29	990
Electrical	3.35	1061
General	3.17	1589
Industrial	3.27	1024
Mechanical	3.13	969
Earth Sciences		
Geology	3.20	984
Petroleum Engineering	3.41	216

*Departments or programs with less than 200 enrollments omitted.

TABLE 3

GRADES EARNED IN FOUR SCHOOL DIVISIONS
BY MAJORS OF EACH DIVISION, 1979-80 DATA
(Means; cases in parentheses)

	Courses in:				
Majors in:	Humanities	Social Sciences	Natural Sciences	Engineering	Total*
Humanities	3.49	3.34	3.27	3.17	3.47
	(5049)	(1051)	(1047)	(1047)	(10250)
Social Sciences	3.39	3.36	3.21	3.25	3.36
	(3340)	(7066)	(1707)	(430)	(15608)
Natural Sciences	3.50	3.53	3.40	3.35	3.46
	(1340)	(896)	(5706)	(255)	(9711)
Interdepartmental Programs	3.42	3.36	3.10	3.28	3.34
	(2253)	(2030)	(1882)	(388)	(10169)
Engineering	3.39	3.38	3.21	3.20	3.26
	(871)	(588)	(2791)	(4947)	(10287)
Earth Sciences	3.39	3.36	2.96	3.17	3.27
	(272)	(149)	(474)	(133)	(2001)
				Total mean	3.37* (58026)
Total	3.45	3.35	3.27	3.21	
	(13131)	(11780)	(13607)	(6394)	
All courses			3.33		
			(44912)		

*Totals in this column include courses outside the four largest school divisions.

TABLE 4

GRADUATES EARNING PHI BETA KAPPA AND UNIVERSITY DISTINCTION, BY SCHOOL DIVISION, 1975-76 TO 1981-82
(Percentages)

School Division	1975-76	1976-77	1977-78	1978-79	1979-80	1980-81	1981-82	Mean
Phi Beta Kappa								
Humanities	-	14	12	13	12	11	9	12
Social Sciences	-	9	9	8	8	8	7	8
Natural Sciences	-	15	15	14	15	13	10	14
Interdepartmental Programs	-	12	13	9	11	8	6	10
Engineering	-	7	6	3	7	4	7	6
Earth Sciences	-	3	5	13	8	5	5	6
All Graduates	-	11	11	10	10	9	8	10
Distinction								
Humanities	24	28	27	25	23	24	23	25
Social Sciences	22	24	24	19	22	20	20	22
Natural Sciences	26	29	22	23	18	21	24	23
Interdepartmental Programs	25	24	23	24	19	21	18	22
Engineering	26	26	19	15	20	24	23	22
Earth Sciences	22	16	2	26	14	25	20	18
All Graduates	24	26	23	22	20	22	22	23

TABLE 5

CORRELATIONS BETWEEN GRADES EARNED AND SCHOLASTIC APTITUDE TEST SCORES BY SCHOOL DIVISION AND MAJOR*
(Means for 1979-80 graduates)

	Correlations between grades and: Verbal SAT Score	Mathematics SAT Score	Enrollments
All Courses	.159	.144	58304
Courses in:			
Humanities	.193	.139	13311
Social Sciences	.221	.193	11839
Natural Sciences	.137	.238	13566
Engineering	.126	.236	6248
Students majoring in:			
Humanities	.155	.139	10131
Social Sciences	.193	.168	15326
Natural Sciences	.086	.166	9612
Interdepartmental Programs	.172	.111	9986
Engineering	.145	.233	9915
Earth Sciences	.081	.153	1982

*Pearson zero-order correlations; all coefficients significant beyond .001 level.

TABLE 6

FACTORS AFFECTING GRADES EARNED, 1979-80
GRADUATES, BY SCHOOL DIVISION
(Multiple regression analyses)

	All Enrollments (N=48096)				
Independent variable	b	Std. error	β	F-ratio	Zero-order r
Verbal SAT	.00088	.00004	.099	414.5	.17
Math SAT	.0013	.00004	.155	895.4	.17
Sex of student	.013	.0065	.009	3.9[ns]	.03
Class size	-.034	.0012	-.139	816.9	-.18
Instructor type	.036	.0062	.027	33.6	.06
Course level	.031	.0063	.023	24.6	.03
Own school division	.027	.0060	.020	20.7	.02
Technology/science school division	.164	.0062	.122	702.7	.12

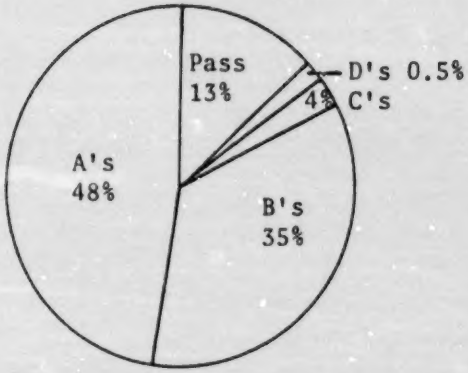
r^2 = .086, constant = 2.24

	Enrollments in:				
	Humanities	Social Sciences	Natural Sciences	Engineering	Earth Sciences
	β	β	β	β	β
Verbal SAT	.162	.163	.042	.020[ns]	.032[ns]
Math SAT	.071	.135	.187	.228	.158
Sex of Student	-.003[ns]	-.003[ns]	.061	.000[ns]	-.070
Class size	-.195	-.129	-.111	-.117	-.256
Instructor type	.010[ns]	.022	.076	.019[ns]	.014[ns]
Course level	.046	-.022	.042	.038	-.057
Own school division	.020	-.048	.105	-.061	-.027[ns]
r^2	.086	.079	.102	.075	.087
Constant	2.67	2.03	2.08	1.72	2.31
Enrollments	12813	11740	13484	6214	1392

ns = not significant at .05 level

FIGURE 1
DISTRIBUTION OF GRADES AND MEAN GRADES BY
SCHOOL DIVISION, 1976-77 AND 1978-79 TO 1981-82
(mean figures)

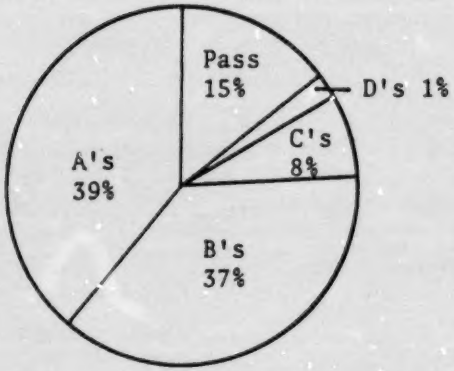
HUMANITIES COURSES



N = 21,174 enrollments

Mean = 3.51

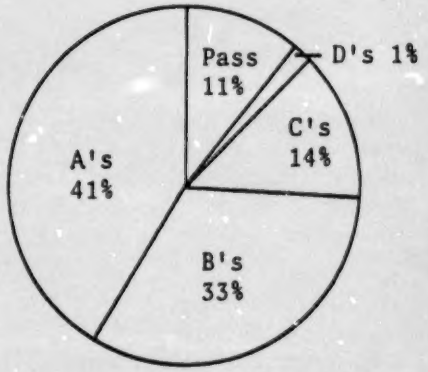
SOCIAL SCIENCE COURSES



N = 15,382 enrollments

Mean = 3.34

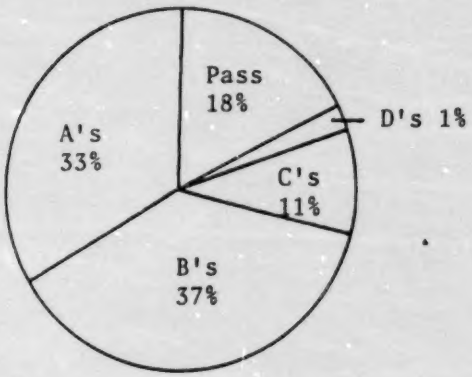
NATURAL SCIENCES COURSES



N = 17,516 enrollments

Mean = 3.28

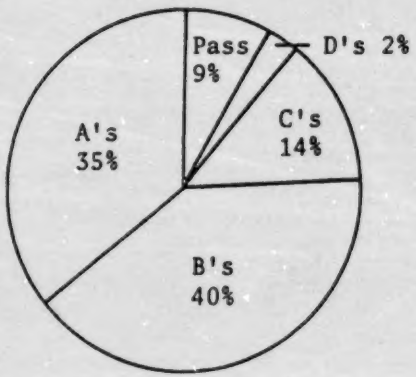
INTERDEPARTMENTAL PROGRAM COURSES



N = 4,010 enrollments

Mean = 3.24

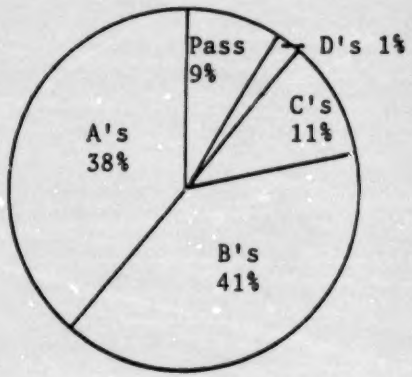
ENGINEERING COURSES



N = 7,222 enrollments

Mean = 3.21

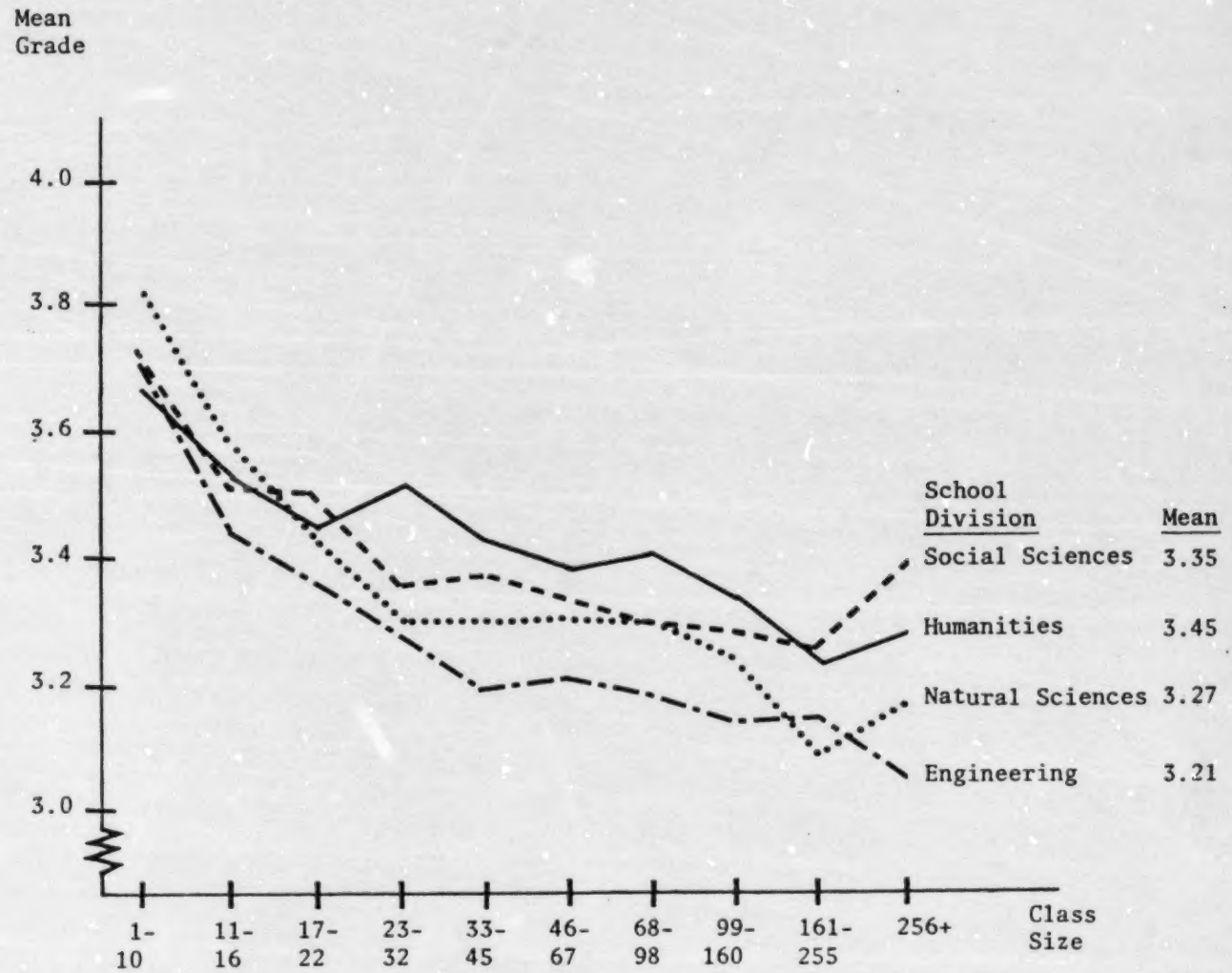
EARTH SCIENCE COURSES



N = 1,530 enrollments

Mean = 3.27

FIGURE 2
MEAN GRADES BY CLASS SIZE DECILES AND
SCHOOL DIVISION, 1979-80 DATA



Enrollments: Humanities 13,142
Social Sciences 11,951
Natural Sciences 13,796
Engineering 6,452